

SEQUENCE LISTING

(1) GENERAL INFORMATION

(i) APPLICANT: LaBrie, Samuel T.  
Lal, Preeti  
Murry, Lynn E.

(ii) TITLE OF THE INVENTION: NOVEL TUBBY HOMOLOGUE

(iii) NUMBER OF SEQUENCES: 4

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: Incyte Pharmaceuticals, Inc.  
(B) STREET: 3174 Porter Dr.  
(C) CITY: Palo Alto  
(D) STATE: CA  
(E) COUNTRY: USA  
(F) ZIP: 94304

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Diskette  
(B) COMPUTER: IBM Compatible  
(C) OPERATING SYSTEM: DOS  
(D) SOFTWARE: FastSEQ for Windows Version 2.0

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER: To Be Assigned  
(B) FILING DATE: Filed Herewith

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER:  
(B) FILING DATE:

(viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: Billings, Lucy J  
(B) REGISTRATION NUMBER: 36,749  
(C) REFERENCE/DOCKET NUMBER: PF-0232US

(ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: 415-855-0555  
(B) TELEFAX: 415-845-4166

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 491 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(vii) IMMEDIATE SOURCE:

(A) LIBRARY: HNT2NOT01  
(B) CLONE: Concensus

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

PF-0232 US

Met	Glu	Ala	Ser	Arg	Cys	Arg	Leu	Ser	Pro	Ser	Gly	Asp	Ser	Val	Phe
1				5					10					15	
His	Glu	Glu	Met	Met	Lys	Met	Arg	Gln	Ala	Lys	Leu	Asp	Tyr	Gln	Arg
			20					25					30		
Leu	Leu	Leu	Glu	Lys	Arg	Gln	Arg	Lys	Lys	Arg	Leu	Glu	Pro	Phe	Met
		35				40						45			
Val	Gln	Pro	Asn	Pro	Glu	Ala	Arg	Leu	Arg	Arg	Ala	Lys	Pro	Arg	Ala
		50				55					60				
Ser	Asp	Glu	Gln	Thr	Pro	Leu	Val	Asn	Cys	His	Thr	Pro	His	Ser	Asn
65					70				75						80
Val	Ile	Leu	His	Gly	Ile	Asp	Gly	Pro	Ala	Ala	Val	Leu	Lys	Pro	Asp
				85				90						95	
Glu	Val	His	Ala	Pro	Ser	Val	Ser	Ser	Ser	Val	Val	Glu	Glu	Asp	Ala
			100					105					110		
Glu	Asn	Thr	Val	Asp	Thr	Ala	Ser	Lys	Pro	Gly	Leu	Gln	Glu	Arg	Leu
		115						120					125		
Gln	Lys	His	Asp	Ile	Ser	Glu	Ser	Val	Asn	Phe	Asp	Glu	Glu	Thr	Asp
		130				135					140				
Gly	Ile	Ser	Gln	Ser	Ala	Cys	Leu	Glu	Arg	Pro	Asn	Ser	Ala	Ser	Ser
145					150					155					160
Gln	Asn	Ser	Thr	Asp	Thr	Gly	Thr	Ser	Gly	Ser	Ala	Thr	Ala	Ala	Gln
				165					170					175	
Pro	Ala	Asp	Asn	Leu	Leu	Gly	Asp	Ile	Asp	Asp	Leu	Glu	Asp	Phe	Val
			180					185					190		
Tyr	Ser	Pro	Ala	Pro	Gln	Gly	Val	Thr	Val	Arg	Cys	Arg	Ile	Ile	Arg
		195				200						205			
Asp	Lys	Arg	Gly	Met	Asp	Arg	Gly	Leu	Phe	Pro	Thr	Tyr	Tyr	Met	Tyr
		210				215					220				
Leu	Glu	Lys	Glu	Glu	Asn	Gln	Lys	Ile	Phe	Leu	Leu	Ala	Ala	Arg	Lys
225					230					235					240
Arg	Lys	Lys	Ser	Lys	Thr	Ala	Asn	Tyr	Leu	Ile	Ser	Ile	Asp	Pro	Val
				245					250					255	
Asp	Leu	Ser	Arg	Glu	Gly	Glu	Ser	Tyr	Val	Gly	Lys	Leu	Arg	Ser	Asn
			260					265					270		
Leu	Met	Gly	Thr	Lys	Phe	Thr	Val	Tyr	Asp	Arg	Gly	Ile	Cys	Pro	Met
		275					280					285			
Lys	Gly	Arg	Gly	Leu	Val	Gly	Ala	Ala	His	Thr	Arg	Gln	Glu	Leu	Ala
		290				295					300				
Ala	Ile	Ser	Tyr	Glu	Thr	Asn	Val	Leu	Gly	Phe	Lys	Gly	Pro	Arg	Lys
305					310					315					320
Met	Ser	Val	Ile	Ile	Pro	Gly	Met	Thr	Leu	Asn	His	Lys	Gln	Ile	Pro
				325					330					335	
Tyr	Gln	Pro	Gln	Asn	Asn	His	Asp	Ser	Leu	Leu	Ser	Arg	Trp	Gln	Asn
			340					345					350		
Arg	Thr	Met	Glu	Asn	Leu	Val	Glu	Leu	His	Asn	Lys	Ala	Pro	Val	Trp
		355					360					365			
Asn	Ser	Asp	Thr	Gln	Ser	Tyr	Val	Leu	Asn	Phe	Arg	Gly	Arg	Val	Thr
		370				375					380				
Gln	Ala	Ser	Val	Lys	Asn	Phe	Gln	Ile	Val	His	Lys	Asn	Asp	Pro	Asp
385					390					395					400
Tyr	Ile	Val	Met	Gln	Phe	Gly	Arg	Val	Ala	Asp	Asp	Val	Phe	Thr	Leu
			405						410					415	
Asp	Tyr	Asn	Tyr	Pro	Leu	Cys	Ala	Val	Gln	Ala	Phe	Gly	Ile	Gly	Leu
			420					425					430		
Ser	Ser	Phe	Asp	Lys	Arg	Ile	Gln	Thr	Leu	Arg	Met	Gln	Glu	Leu	Cys
		435					440					445			
Glu	Leu	His	Arg	Gln	His	His	Ser	Ala	Ala	Ser	Leu	Val	His	Arg	Thr
		450				455					460				
Ala	Cys	Gln	Arg	Trp	Val	Gly	His	Pro	Trp	Arg	Gln	Leu	Pro	Gln	Ser
465					470					475					480
Ser	Leu	Val	Gly	Pro	Asp	Leu	Xaa	Leu	His	Met					

485

490

## (2) INFORMATION FOR SEQ ID NO:2:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1525 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: cDNA

## (vii) IMMEDIATE SOURCE:

- (A) LIBRARY: HNT2NOT01
- (B) CLONE: Consensus

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

GCACGAGGTG	GGGGCTTTCC	TCGGTGGCGG	GCATGGAGGC	TTGCGGCTGC	CGGCTCAGTC	60
CCAGCGGCGA	CAGTGTCTTC	CATGAAGAAA	TGATGAAGAT	GCGACAGGCT	AAGCTGGATT	120
ATCAGAGGCT	ACTACTTGAG	AAGAGGCCAA	GGAAAAAGCG	CCTTGAGCCA	TTTATGGTGC	180
AGCCCAATCC	AGAAGCCAGG	CTACGTCGGG	CAAAGCCAAAG	GGCCAGTGAT	GAGCAGACTC	240
CCTTGGTGAA	CTGTCATACT	CCCCACAGCA	ATGTCATCTT	ACATGGTATT	GATGGTCCAG	300
CTGCTGTCCT	GAAACCAGAC	GAAGTTCATG	CTCCATCAGT	AAGCTCCTCT	GTTGTGGAAG	360
AAGATGCTGA	AAACACCGTG	GATACTGCTT	CCAAGCCAGG	ACTTCAGGAG	CGTCTCCAAA	420
AGCATGATAT	CTCTGAAAGT	GTGAACTTCG	ATGAGGAGAC	TGATGGAATA	TCCCAGTCAG	480
CATGTTTAGA	AAGACCCAAT	TCTGCATCAA	GCCAGAATTC	AACCGATACA	GGCACTTCCG	540
GTTCTGCTAC	TGCCGCCCAA	CCAGCTGATA	ACCTCCTGGG	AGACATAGAC	GACCTGGAGG	600
ACTTTGTGTA	TAGTCCTGCC	CCTCAAGGTG	TCACAGTAAG	ATGTCGGATA	ATCCGGGATA	660
AAAGGGGAAT	GGATCGGGGT	CTCTTCCCCA	CCTACTATAT	GTAATTGGAA	AAAGAAGAAA	720
ATCAGAAGAT	ATTTCTTCTT	GCAGCTAGAA	AGCGGAAAAA	GAGCAAAACA	GCCAACTACC	780
TTATCTCCAT	TGATCCAGTT	GATTTATCTC	GTGAAGGAGA	AAGTTATGTC	GGCAAGCTTA	840
GATCCAACCT	CATGGGGACC	AAGTTTACAG	TTTATGACCG	TGGCATCTGC	CCCATGAAGG	900
GCCGGGGTTT	GGTAGGAGCG	GCCCACACCC	GGCAGGAGCT	GGCTGCCATC	TCCTATGAAA	960
CAAACGTACT	TGGATTAAAA	GGTCCTAGGA	AAATGTCTGT	GATCATTCCT	GGAATGACAC	1020
TGAATCATAA	GCAGATCCCC	TATCAGCCAC	AAAACAACCA	TGACAGTTTG	CTCTCAAGGT	1080
GGCAGAACAG	AACTATGGAA	AATCTGGTTG	AGCTGCACAA	CAAGGCCCCC	GTCTGGAACA	1140
GTGACACTCA	GTCTATGTC	CTCAACTTCC	GTGGCCGGGT	CACTCAGGCG	TCTGTGAAGA	1200
ACTTCCAGAT	AGTCCACAAA	AATGACCCTG	ATTATATAGT	CATGCAGTTT	GGACGTGTGG	1260
CAGATGACGT	GTTCACTACTG	GATTACAAC	ACCCACTTTG	TGCAGTACAG	GCCTTTGGCA	1320
TCGGTCTTTC	TAGCTTTGAC	AAACGTATCC	AAACCTTGAG	AATGCAGGAG	CTCTGTGAGC	1380
TCCACCGTCA	GCACCATTCA	GCTGCATCCC	TTGTGCACAG	GACTGCCTGC	CAGCGTTGGG	1440
TGGGACACCC	GTGGCGGCAG	CTCCCTCAGT	CTTCCCTTGT	CGGCCCTGAC	CTNTNACTAC	1500
ATATGTAGNA	GCCCAGAGACC	AAAAA				1525

## (2) INFORMATION FOR SEQ ID NO:3:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 505 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: peptide

## (vii) IMMEDIATE SOURCE:

- (A) LIBRARY: GenBank
- (B) CLONE: 1279766

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

Met	Thr	Ser	Lys	Pro	His	Ser	Asp	Trp	Ile	Pro	Tyr	Ser	Val	Leu	Asp
1				5					10					15	
Asp	Glu	Gly	Ser	Asn	Leu	Arg	Gln	Gln	Lys	Leu	Asp	Arg	Gln	Arg	Ala
			20					25					30		
Leu	Leu	Glu	Gln	Lys	Gln	Lys	Lys	Lys	Arg	Gln	Glu	Pro	Leu	Met	Val
		35					40					45			
Gln	Ala	Asn	Ala	Asp	Gly	Arg	Pro	Arg	Ser	Arg	Arg	Ala	Arg	Gln	Ser
	50					55					60				
Glu	Glu	Gln	Ala	Pro	Leu	Val	Glu	Ser	Tyr	Leu	Ser	Ser	Ser	Gly	Ser
65					70					75					80
Thr	Ser	Tyr	Gln	Val	Gln	Glu	Ala	Asp	Ser	Ile	Ala	Ser	Val	Gln	Leu
			85						90					95	
Gly	Ala	Thr	Arg	Pro	Pro	Ala	Pro	Ala	Ser	Ala	Lys	Lys	Ser	Lys	Gly
			100					105						110	
Ala	Ala	Ala	Ser	Gly	Gly	Gln	Gly	Gly	Ala	Pro	Arg	Lys	Glu	Lys	Lys
		115						120					125		
Gly	Lys	His	Lys	Gly	Thr	Ser	Gly	Pro	Ala	Thr	Leu	Ala	Glu	Asp	Lys
	130					135					140				
Ser	Glu	Ala	Gln	Gly	Pro	Val	Gln	Ile	Leu	Thr	Val	Gly	Gln	Ser	Asp
145					150					155					160
His	Asp	Lys	Asp	Ala	Gly	Glu	Thr	Ala	Ala	Gly	Gly	Gly	Ala	Gln	Pro
			165						170					175	
Ser	Gly	Gln	Asp	Leu	Arg	Ala	Thr	Met	Gln	Arg	Lys	Gly	Ile	Ser	Ser
			180					185					190		
Ser	Met	Ser	Phe	Asp	Glu	Asp	Glu	Asp	Glu	Asp	Glu	Asn	Ser	Ser	Ser
		195					200					205			
Ser	Ser	Gln	Leu	Asn	Ser	Asn	Thr	Arg	Pro	Ser	Ser	Ala	Thr	Ser	Arg
		210				215					220				
Lys	Ser	Ile	Arg	Glu	Ala	Ala	Ser	Ala	Pro	Ser	Pro	Ala	Ala	Pro	Glu
225				230						235					240
Pro	Pro	Val	Asp	Ile	Glu	Val	Gln	Asp	Leu	Glu	Glu	Phe	Ala	Leu	Arg
			245						250					255	
Pro	Ala	Pro	Gln	Gly	Ile	Thr	Ile	Lys	Cys	Arg	Ile	Thr	Arg	Asp	Lys
		260						265					270		
Lys	Gly	Met	Asp	Arg	Gly	Met	Tyr	Pro	Thr	Tyr	Phe	Leu	His	Leu	Asp
		275					280					285			
Arg	Glu	Asp	Gly	Lys	Lys	Val	Phe	Leu	Leu	Ala	Gly	Arg	Lys	Arg	Lys
		290				295					300				
Lys	Ser	Lys	Thr	Ser	Asn	Tyr	Leu	Ile	Ser	Val	Asp	Pro	Thr	Asp	Leu
305				310						315					320
Ser	Arg	Gly	Gly	Asp	Ser	Tyr	Ile	Gly	Lys	Leu	Arg	Ser	Asn	Leu	Met
			325						330					335	
Gly	Thr	Lys	Phe	Thr	Val	Tyr	Asp	Asn	Gly	Val	Asn	Pro	Gln	Lys	Ala
		340						345					350		
Ser	Ser	Ser	Thr	Leu	Glu	Ser	Gly	Thr	Leu	Arg	Gln	Glu	Leu	Ala	Ala
		355					360					365			
Val	Cys	Tyr	Glu	Thr	Asn	Val	Leu	Gly	Phe	Lys	Gly	Pro	Arg	Lys	Met
		370				375					380				
Ser	Val	Ile	Val	Pro	Gly	Met	Asn	Met	Val	His	Glu	Arg	Val	Cys	Ile
385				390						395					400
Arg	Pro	Arg	Asn	Glu	His	Glu	Thr	Leu	Leu	Ala	Arg	Trp	Gln	Asn	Lys
			405						410					415	
Asn	Thr	Glu	Ser	Ile	Ile	Glu	Leu	Gln	Asn	Lys	Thr	Pro	Val	Trp	Asn
		420						425					430		
Asp	Asp	Thr	Gln	Ser	Tyr	Val	Leu	Asn	Phe	His	Gly	Arg	Val	Thr	Gln
		435					440					445			
Ala	Ser	Val	Lys	Asn	Phe	Gln	Ile	Ile	His	Gly	Asn	Asp	Pro	Asp	Tyr
		450				455					460				
Ile	Val	Met	Gln	Phe	Gly	Arg	Val	Ala	Glu	Asp	Val	Phe	Thr	Met	Asp
465				470						475					480
Tyr	Asn	Tyr	Pro	Leu	Cys	Ala	Leu	Gln	Ala	Phe	Ala	Ile	Ala	Leu	Ser

485 490 495  
 Ser Phe Asp Ser Lys Leu Ala Cys Glu  
 500 505

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 506 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(vii) IMMEDIATE SOURCE:

- (A) LIBRARY: GenBank  
(B) CLONE: 1305497

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Met 1	Thr	Ser	Lys	Pro 5	His	Ser	Asp	Trp	Ile 10	Pro	Tyr	Ser	Val	Leu 15	Asp
Asp	Glu	Gly	Arg 20	Asn	Leu	Arg	Gln	Gln 25	Lys	Leu	Asp	Arg	Gln 30	Arg	Ala
Leu	Leu	Glu 35	Gln	Lys	Gln	Lys	Lys 40	Lys	Arg	Gln	Glu	Pro 45	Leu	Met	Val
Gln	Ala 50	Asn	Ala	Asp	Gly	Arg 55	Pro	Arg	Ser	Arg	Arg 60	Ala	Arg	Gln	Ser
Glu 65	Glu	Gln	Ala	Pro	Leu 70	Val	Glu	Ser	Tyr	Leu 75	Ser	Ser	Ser	Gly	Ser 80
Thr	Ser	Tyr	Gln	Val 85	Gln	Glu	Ala	Asp	Ser 90	Leu	Ala	Ser	Val	Gln 95	Leu
Gly	Ala	Thr 100	Arg	Pro	Thr	Ala	Pro	Ala 105	Ser	Ala	Lys	Arg	Thr 110	Lys	Ala
Ala	Ala	Thr 115	Ala	Gly	Gly	Gln	Gly 120	Gly	Ala	Ala	Arg	Lys 125	Glu	Lys	Lys
Gly	Lys 130	His	Lys	Gly	Thr	Ser 135	Gly	Pro	Ala	Ala	Leu 140	Ala	Glu	Asp	Lys
Ser 145	Glu	Ala	Gln	Gly	Pro 150	Val	Gln	Ile	Leu	Thr 155	Val	Gly	Gln	Ser	Asp 160
His	Ala	Gln	Asp	Ala 165	Gly	Glu	Thr	Ala	Ala 170	Gly	Gly	Gly	Glu	Arg 175	Pro
Ser	Gly	Gln	Asp 180	Leu	Arg	Ala	Thr	Met 185	Gln	Arg	Lys	Gly	Ile 190	Ser	Ser
Ser	Met 195	Ser	Phe	Asp	Glu	Asp	Glu 200	Glu	Asp	Glu	Glu	Glu 205	Asn	Ser	Ser
Ser	Ser 210	Ser	Gln	Leu	Asn	Ser 215	Asn	Thr	Arg	Pro	Ser 220	Ser	Ala	Thr	Ser
Arg 225	Lys	Ser	Val	Arg 230	Glu	Ala	Ala	Ser	Ala	Pro 235	Ser	Pro	Thr	Ala	Pro 240
Glu	Gln	Pro	Val	Asp 245	Val	Glu	Val	Gln	Asp 250	Leu	Glu	Glu	Phe	Ala 255	Leu
Arg	Pro	Ala	Pro 260	Gln	Gly	Ile	Thr	Ile 265	Lys	Cys	Arg	Ile	Thr 270	Arg	Asp
Lys	Lys	Gly 275	Met	Asp	Arg	Gly	Met 280	Tyr	Pro	Thr	Tyr	Phe 285	Leu	His	Leu
Asp	Arg 290	Glu	Asp	Gly	Lys	Lys 295	Val	Phe	Leu	Leu	Ala 300	Gly	Arg	Lys	Arg
Lys 305	Lys	Ser	Lys	Thr	Ser 310	Asn	Tyr	Leu	Ile	Ser 315	Val	Asp	Pro	Thr	Asp 320
Leu	Ser	Arg	Gly	Gly	Asp	Ser	Tyr	Ile	Gly	Lys	Leu	Arg	Ser	Asn	Leu

Met	Gly	Thr	Lys	325	Phe	Thr	Val	Tyr	Asp	330	Asn	Gly	Val	Asn	Pro	335	Gln	Lys
			340						345						350			
Ala	Ser	Ser	Ser	Thr	Leu	Glu	Ser	Gly	Thr	Leu	Arg	Gln	Glu	Leu	Ala			
		355					360					365						
Ala	Val	Cys	Tyr	Glu	Thr	Asn	Val	Leu	Gly	Phe	Lys	Gly	Pro	Arg	Lys			
	370					375					380							
Met	Ser	Val	Ile	Val	Pro	Gly	Met	Asn	Met	Val	His	Glu	Arg	Val	Ser			
385					390					395					400			
Ile	Arg	Pro	Arg	Asn	Glu	His	Glu	Thr	Leu	Leu	Ala	Arg	Trp	Gln	Asn			
			405					410						415				
Lys	Asn	Thr	Glu	Ser	Ile	Ile	Glu	Leu	Gln	Asn	Lys	Thr	Pro	Val	Trp			
		420					425						430					
Asn	Asp	Asp	Thr	Gln	Ser	Tyr	Val	Leu	Asn	Phe	His	Gly	Arg	Val	Thr			
		435				440						445						
Gln	Ala	Ser	Val	Lys	Asn	Phe	Gln	Ile	Ile	His	Gly	Asn	Asp	Pro	Asp			
	450				455					460								
Tyr	Ile	Val	Met	Gln	Phe	Gly	Arg	Val	Ala	Glu	Asp	Val	Phe	Thr	Met			
465				470					475					480				
Asp	Tyr	Asn	Tyr	Pro	Leu	Cys	Ala	Leu	Gln	Ala	Phe	Ala	Ile	Ala	Leu			
			485					490						495				
Ser	Ser	Phe	Asp	Ser	Lys	Leu	Ala	Cys	Glu									
			500					505										